CROP HEIGHT ESTIMATION OF WHEAT USING SENTINEL-1 SATELLITE IMAGERY: PRELIMINARY RESULTS

O. G. Narin¹, C. Bayik², A. Sekertekin³, S. Madenoglu⁴, M. O. Pinar⁵, S. Abdikan⁶, F. Balik Sanli⁷

¹Department of Geomatics Engineering, Afyon Kocatepe University, Afyonkarahisar, Türkiye - gokberknarin@aku.edu.tr
 ²Department of Geomatics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Türkiye - caglarbayik@hotmail.com
 ³Department of Architecture and Town Planning, Igdir University, Igdir, Türkiye – aliihsan.sekertekin@igdir.edu.tr
 ⁴Ministry of Agriculture and Forestry, General Directorate of Agricultural Research and Policies (TAGEM), Ankara, Türkiye - sevinc.madenoglu@tarimorman.gov.tr

⁵ Ministry of Agriculture and Forestry, TAGEM, Transitional Zone Agricultural Research Institute, Eskisehir, Turkey melisozge.pinar@tarimorman.gov.tr

⁶Department of Geomatics Engineering, Hacettepe University, Ankara, Türkiye - sayginabdikan@hacettepe.edu.tr ⁷Department of Geomatics Engineering, Yildiz Technical University, Istanbul, Türkiye - fbalik@yildiz.edu.tr

KEY WORDS: Remote Sensing, Crop Height, Sentinel-1, Wheat, Random Forest, Linear Regression.

ABSTRACT:

Wheat is one of the primary crop productions for half of the global production. It is the most exported cereal, which reached up to 40% according to the records of the Food and Agriculture Organization of the United Nations (FAO) in 2020. The relationship between the crop parameters and remote sensing tools has become essential for the temporal monitoring and estimations of crop growth. In this study, we investigate the relationship between crop height and Synthetic Aperture Radar (SAR) backscatter. Plant height was collected twice, in the early (13 April 2023) and late stages (21 June 2023) of wheat, for a total of 70 samples. Sentinel-1 SAR data was also attempted to be synchronized with ground measurements. For this purpose, 8 images were obtained, 4 in both ascending (8 and 15 April, 16 and 26 June) and descending (9 and 14 April, 20 and 25 June) directions. The basic steps of image pre-processing, calibration, filtering, and topography correction, were applied to each image. By evaluating the correlation coefficients between plant height and images, data with low correlations were excluded and the first three data, dated 8 April and 19 June, were used for plant height estimation. For prediction purposes, Linear Regression (LR) and Random Forest (RF) methods were evaluated with three different training data sets. The highest correlation and minimum error value were achieved by VH polarization with random forest (r = 0.868, RMSE = 5.377 cm) in the early stage and LR (r = 0.616, RMSE = 14.451 cm) in the late stage. In general, higher training (80%) and lower test data (20%) produced better results.

1. INTRODUCTION

Crops are crucial nutritional sources that contribute to the sustainability of society and human life (Li et al., 2022). Effective crop management is required throughout all phenological stages of the crops to increase the yield. Crop height, one of the key biophysical parameters, provides valuable information about crop growth, and it is an important factor used in many agricultural activities such as crop health assessment, phenology tracking, biomass and yield estimation, and precision fertilization (Erten et al., 2016; Xie et al., 2021). Thus, accurate, reliable, and systematic monitoring and retrieval of crop height is essential to support agricultural crop management services (Romero-Puig and Lopez-Sanchez, 2021).

Demand for food is increasing globally due to population growth, urbanization, and changing eating habits. Wheat plays an important role in meeting this demand. Traditional methods of crop growth monitoring using field surveys and ground sensors are labor-intensive and insufficient to provide timely and comprehensive information on large-scale agricultural environments.

In response to these challenges, Remote sensing technology enables monitoring of the earth's surface on a large scale with varying spatial and temporal resolutions. Unlike the traditional measurement methods of crop height, remote sensing helps to estimate time- and cost-effective crop height using optical and Synthetic Aperture Radar (SAR) sensors (Lee et al., 2018;

Yuzugullu et al., 2018; Nasirzadehdizaji et al., 2019; Abdikan et al., 2023). Among them, the Sentinel 1 satellite equipped with a synthetic aperture radar is characterized by the ability to collect data regardless of weather and lighting conditions. Concerning the literature, many researchers have focused on the retrieval of the crop height using Sentinel-1 SAR data. Yang et al. (2022) utilized time series Sentinel 1 images to estimate the spatiotemporal distribution of rice height. They stated that the estimation of the rice height by particle filter gave a Root Mean Square Error (RMSE) of 7.36 cm and a coefficient of determination (R²) of 0.95 compared to the simplified water cloud model (RMSE = 12.59 cm and R2 = 0.86). Kaplan et al. (2023) used Sentinel-1 and Sentinel-2 imagery for estimating cotton crop coefficient, height, and Leaf Area Index (LAI). Their results revealed that the most robust Sentinel-1 models were obtained by applying an innovative local incidence angle normalization method with $R^2 = 0.6586$, and RMSE = 18 cm for the height estimation. Harfenmeister et al. (2019) analyzed temporal and spatial characteristics of crop parameters using Sentinel-1 backscatter data. They expressed that some fields achieve R² values greater than 0.9 for VV backscatter and multiple regression on plant height, with RMSE values around 10 cm. Ndikumana et al. (2018) aimed to study the capabilities of multitemporal radar images for rice height and dry biomass retrievals using Sentinel-1 data. The rice height estimation accuracy showed that the rice height retrieval was highly correlated with the field rice height from dual-polarization, in which the random forest presented the best performance with $R^2=0.92$ and the RMSE 16% (7.9 cm). Sharifi and

Hosseingholizadeh used Sentinel-1 data to estimate the height and biomass of rice crops in Astaneh-ye Ashrafiyeh, Iran. They declared that the efficiency of the nonparametric methods (SVR and RVR) is much better than that of the parametric regression (MLR) for rice parameter estimations. Their results showed that the RVR provided better performance than the MLR and SVR with R^2 =0.92 and RMSE=10.9. Nasirzadehdizaji et al. (2019) showed the contribution of coherence data derived from interferometric processes in addition to backscatter value for crop monitoring.

As presented above, many studies researched crop height estimation using Sentinel-1 SAR data; however, few studies considered the estimation of wheat crop height. Therefore, this study aims to investigate the potential of the dual-polarized Sentinel-1 backscatter data for wheat height estimation on preliminary results of the field measurements. The selected wheat fields are located in the capital city of Türkiye, Ankara.

2. METHOD

2.1 Study Area

To estimate plant height, fieldwork was carried out on 2 different dates (13 April and 21 June) in the agricultural area within the borders of Ankara province, in Central Anatolia, Turkiye. According to the Köppen climate classification, the study area in Ankara is classified as a cold semi-arid climate (Hot-summer Mediterranean climate) (Yılmaz and Çiçek 2018).



Figure 1. Study area. Frames show the coverage of the satellite images. Black and green are the ascending, and blue with yellow are descending mode image acquisitions. The red circle shape is the location of the study area in Ankara District.

In this region where the continental climate prevails, winter temperatures are low and summers are hot. The average temperature value of 25 years in Ankara is 12 °C annually. The average temperature drops to 0°C in winter and reaches 23.5°C in July and August in summer. The average monthly total precipitation amount (mm) is minimum in summer (12-14 mm) and reaches its maximum value (~45 mm) in winter (MGM 2023). The altitude in the region reaches up to 1000 m. In general, in central Anatolia, sowing starts end of September and begins of October. The earing period starts mid to end of May. The harvesting period is between the two weeks of July. According to Biologische Bundesanstalt, Bundessortenamt und CHemische (BBCH) scale the fields are in the stem elongation and development of fruit periods during the data collection in April and June (Meier 2018).

2.2 Sentinel-1 Data

Sentinel-1 satellite images, which are available to users free of charge, were used in the study. Since the Sentinel-1B satellite has not been able to provide data since December 2021, only images from the Sentinel-1A satellite were acquired. The image acquisition was planned to be taken simultaneously with the field studies carried out in April and June. For both descending and ascending orbits, the closest images were taken at the time of ground measurements. A total of 8 Sentinel-1 GRD (Ground Range Detected) images were used (Figure 1). In the analysis, processes were carried out in 4 different frames; 2 different ascending and 2 different descending (Table 1).

Table 1. Characteristics of Sentinel-1A over the study are

Orbit Number	Data	Satellite Pass	Incidence Angle (°)
87	15 April	Ascending	39.04
	26 June		
160	08 April	Ascending	38.93
	19 June		
65	14 April	Descending	39.08
	25 June		
167	09 April	Descending	39.11
	20 June		

Each image has dual polarization: vertical-vertical (VV), and vertical-horizontal (VH). Open-source tools of Sentinel Application Platform (SNAP) (SNAP, 2023) software were used to obtain backscatter coefficients (σ°) using GRD images at 10 m resolution. Four processing steps were applied: i) Apply Orbit File; ii) Radiometric Calibration; iii) Speckle Filter; iv) Range Doppler Terrain Correction using SRTM 1Sec DEM (Digital Elevation Model) data (Figure 2).

The backscatter coefficient (σ°) values were converted to decibel (dB) values using the equation (1);

$$\sigma_{dB} = 10 \cdot \log_{10} \left(\sigma^{\circ} \right) \tag{1}$$

2.3 Regression methods

In our study, two different regression methods Random Forest (RF) and Linear regression (LR) were used to estimate the height of the wheat plant. LR is the most commonly used regression model that gives a linear relationship between two data groups (Equation 2).

$$y = ax + b \tag{2}$$

where y is the dependent variable, x is the independent variable, a is the constant, and b is the coefficient.



Figure 2. Sentinel-1 GRD pre-processing workflow.

RF is an ensemble learning algorithm that stands out as an effective regression method in the field of machine learning. This algorithm creates a powerful model by combining multiple decision trees (Figure 3). Each tree is trained with randomly selected subsamples and features. RF is a reliable and powerful option for regression problems. (Breiman 2001; Rodriguez-Galiano et al., 2015).



Figure 3. Architecture of the random forest model.

The RF method was implemented in the R programming language via the Caret package (Kuhn 2008, R Core Team 2020). When creating the RF model, ntree = 1000, maxnodes = 25 were selected.

2.4 Methodology

Height data were collected from 5 different samples from 70 different points at 20 m intervals from 2 different wheat parcels (Figure 1). The average of 5 height data collected from each point was taken. The relationship between the collected height data of the wheat plant and the dual-polarized Sentinel-1 data was examined. (Table 1). Firstly, the most recent Sentinel-1 images before and after the fieldwork were acquired from the Copernicus hub.

After obtaining backscatter values from Sentinel-1 data, Pearson Correlation (r) values were examined for all images (Table 2). 3 data with r values higher than 0.60 were selected and a model was produced for prediction.

In regression models, training and test data were examined in 3 different scenarios. These are determined as 60% - 40%, 70% - 30%, and 80% - 20% as training and testing, respectively. Ground sample data is randomly divided into training and testing. Wheat height data and local data obtained from Sentinel-1 data were examined according to (RMSE) (2) and Pearson Correlation (r) (3) metrics.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i^{act} - y_i^{cal})}$$
(2)

$$r = \frac{\sum (y_i^{act} - \overline{y_i^{act}})(y_i^{cal} - \overline{y_i^{cal}})}{\sqrt{\sum (y_i^{act} - \overline{y_i^{act}})^2 (y_i^{cal} - \overline{y_i^{cal}})^2}}$$
(3)

where *n* is the number of samples, y_i^{act} is the measured wheat height, y_i^{cal} is the estimation wheat height, $\overline{y_i^{act}}$ and $\overline{y_i^{cal}}$ the average measured and estimation wheat height.

3. RESULTS

Within the scope of the study, the behaviour of both polarization data in two different growth stages of the plant was examined. Then, the relationship between plant height measurements taken from the places shown in Figure 1 and the backscatter value of these locations was examined.

3.1 Backscatter analysis

When the measurements for April are examined, wheat heights vary between 7 cm and 37 cm (Figure 4a, 4b). It seems that the backscatter values show similar behaviour in the ascending Sentinel-1 images taken on two dates (08 and 15 April) close to the ground measurement date (13 April). Considering VH polarization, it is observed that there is an increase ranging up to ~5dB in one week. In VV polarization, the values are closer to each other on both dates and show a higher backscatter value compared to the VH value (Figure 4a).

Images dated 09 April and 14 April were used in the descending satellite image of April. VV polarized data provides higher backscatter than VH polarized data (Figure 4b).

In the data dated April 14, which is closer to the fieldwork date (April 13), both polarimetric data presented closer results to each other in each parcel. There is fluctuation in the results of other dated data. When the measurements for June are examined,

wheat heights vary between 53 cm and 131 cm (Figure 4c, 4d). It seems that the backscatter values show similar behaviour in the ascending Sentinel-1 images taken on two dates (19 and 26 June) close to the ground measurement date (21 June).



ascending June (d) descending June

Images dated June 20 and June 25 were used in the descending satellite image for June. Similar to the previous three datasets, VH results provide lower backscatter than VV results (Figure 4d). In the VV polarization results, while the results of the two dates are closer to each other in the twilight image (Figure 4c), parallel values that do not overlap each other are obtained in the ascending images (Figure 4d).

3.2 Crop height analysis

When examining the correlation between Sentinel-1 data and ground measurements, Ascending data has r values ranging between 0.773 as maximum and 0.029 as minimum. VH polarimetry of April data gave the best correlation (r=0.773). VV polarization comes in second place with a correlation of 0.622. No connection could be established with the image taken on April 15. A correlation of 0.607 was achieved with the VH polarization of the data dated June 19, which is the first image of June. VV polarization data of the same dated data comes in second place with 0.533 (Table 2).

A high correlation could not be established between images taken in the descending orbit and terrestrial data. The highest value, 0.391, was obtained with the VH polarization dated June 14 (Table 2).

Table 2. Correlation (r) values for 70 data. Dark colors show the three best performing correlations.

Ascending							
8 A	8 April 15 April 19 June		April 15 April 19 Jur		lune	26 J	une
VH	VV	VH	VV	VH	VV	VH	VV
0.773	0.622	0.085	0.029	0.607	0.533	0.465	0.353
Descending							
09 April 14 April 2		20 June		25 June			
VH	VV	VH	VV	VH	VV	VH	VV

According to the correlation (r) results values higher than 0.6 were selected for the modelling. The models were produced with VH and VV polarization in the ascending direction in the image dated April 8, and with only VH polarization in the ascending direction on June 19. After this step, descending data were not evaluated because they showed a very low correlation. Correlation (r) results of the test data are given in Table 3 and RMSE results are given in Table 4.

Linear regression and random forest methods were used in the prediction analysis. The data set was divided into three different training and test data (60%/40%, 70%/30%, 80%/20%) and their performances were compared. In April, correlation values varied between 0.868 and 0.731. In general, it is seen that as the training rate increases, the correlation value also increases. Among the results, the result obtained with the 80%/20% ratio provided a higher correlation. It is seen that the RF method gives more successful results as the number of data used in training increases. When the results are examined according to the correlation in Table 3, the best result was found to be 0.868 with the RF method in VH polarization on April 8.

Contrary to this, VH polarized data provided a lower correlation in June than in April. The highest correlation of 0.616 was obtained with low education data using the LR method (Table 3). The LR result obtained with the 70%/30% data set is slightly better than the result obtained with the 60%/40% data set. The result obtained with the 80%/20% data set was the lowest in both LR and RF results.

When the results of RMSE are examined, it is seen that as the number of training decreases in LR, the accuracy decreases, except for June 19 of the 80-20% group (Table 4).

Table 3. Estimation results according to Pearson Correlation (r)

Data Datia 0/		Ascending			
(Train - Test)	Methods	8 April		19 June	
		VH	VV	VH	
60 - 40	Linear Regression	0.749	0.773	0.611	
	Random Forest	0.731	0.737	0.352	
70 - 30	Linear Regression	0.735	0.806	0.616	
	Random Forest	0.746	0.749	0.282	
80 - 20	Linear Regression	0.771	0.854	0.296	
	Random Forest	0.868	0.852	0.191	

It is seen that the accuracy increases with the increase in the training set for the RF method. When the results are examined according to RMSE in Table 4, VH polarization gave the best result with 5.377 cm on April 8. On June 19, the best result was found to be 14.451 cm with the LR method.

Table 4. Estimation results according to RMSE (cm)

		Ascending			
Data Ratio % (Train - Test)	Method	8 April		19 June	
		VH	VV	VH	
60 - 40	Linear Regression	5.774	5.923	16.972	
	Random Forest	5.984	5.763	19.775	
70 - 30	Linear Regression	6.371	6.653	17.884	
	Random Forest	6.135	5.976	21.925	
80 - 20	Linear Regression	6.686	7.436	14.451	
	Random Forest	5.377	5.600	15.744	

4. DISCUSSION AND CONCLUSION

Within the scope of the study, stem elongation and development of fruit periods of the wheat plant, which have two stages according to the BBCH scale, were evaluated using Sentinel-1 satellite images.

When the backscatter values of the SAR data are evaluated, the ascending data (figure 4a) in the first phase of the plant in April provides lower VH backscatter than the descending data (figure 4c, 4d). In the Ascending results, April VV and VH values gave closer results than each other. In both orbital acquisitions, plant backscatter is lower in VH polarization in both phases. Similar results have been found in studies on wheat in the literature. Saad El Imanni et al., (2022) examined wheat phenology in their study in Morocco. According to the Sentinel-1 time series results, backscatter values vary between -15 dB and -25 dB in VH polarization, while they are higher in VV polarization and vary between -8 dB and -16 dB. In their study with Sentinel-1 satellite data, Mercier et al (2020) found that wheat plants gave lower backscatter in VH data. Nasirzadehdizaji et al (2019) compared images taken in ascending and descending directions for different plant species and stated that the VH value was lower than VV in the backscatter values obtained in both acquisition directions for wheat. They also demonstrated that earlier stages of the crops indicated higher backscattering values than later stages which is similar to our results.

In the study conducted by Liao et al (2018) with full polarimetric C-band Radarsat-2 data, it was stated that cross-polarization

(HV) provided a higher correlation when the wheat height varied between 25 and 65 cm, and was much lower when the plant height was greater than 65 cm.

When the backscatter values of the two phases are compared, the values are closer to each other in the stem elongation period, while they are more variable in the second phase. When the plant heights of the two periods are compared, the difference between plant heights on a parcel basis is less in the first period.

The highest correlation was obtained in the stem elongation stage, which is the first stage of the plant. When examining the correlation between Sentinel-1 data and ground measurements, Ascending data gave better results than Descending data in terms of satellite orbit direction (Table 2). When examined in terms of polarization, VH gave better results than VV (Table 2). In addition, the results were better on the dates when data were collected before the fieldwork. The results for both polarizations were better on April 8 than on June 19 (Table 3).

Considering the RMSE values, the error rate increases as plant height increases. However, in proportion to plant height, the error rate obtained in the late season is lower than the error rate obtained in the early season.

In this study, two dates of dual polarimetric Sentinel-1 data analyses based on radar backscattering have been studied for the wheat crop. The investigation is composed of field measurements of crop height collected through field works and its relationship with SAR data. the contribution of a longer time series analysis of SAR data to estimate the crop growth stages could be more beneficial for more accurate analysis (Nasirzadehdizaji et al 2019).

For further analysis, additional features of SAR data i.e. coherence and decomposition will be tested over the study area including additional season of ground measurements data. Optical images can also be used as complementary data.

REFERENCES

Abdikan, S., Sekertekin, A., Narin, O. G., Delen, A., Sanli, F. B., 2023: A comparative analysis of SLR, MLR, ANN, XGBoost and CNN for crop height estimation of sunflower using Sentinel-1 and Sentinel-2. *Advances in Space Research* 71(7), 3045-3059.

Breiman, L. (2001). Random forests. Machine learning, 45, 5-32.

Erten, E., Lopez-Sanchez, J. M., Yuzugullu, O., Hajnsek, I., 2016: Retrieval of agricultural crop height from space: A comparison of SAR techniques. *Remote Sensing of Environment* 187, 130-144.

Harfenmeister, K., Spengler, D., & Weltzien, C., 2019: Analyzing temporal and spatial characteristics of crop parameters using Sentinel-1 backscatter data. *Remote Sensing*, 11(13), 1569.

Kaplan, G., Fine, L., Lukyanov, V., Malachy, N., Tanny, J., & Rozenstein, O., 2023: Using Sentinel-1 and Sentinel-2 imagery for estimating cotton crop coefficient, height, and Leaf Area Index. *Agricultural Water Management*, 276, 108056.

Kuhn, M. 2008. Building predictive models in R using the caret package. *Journal of statistical software*, 28, 1-26.

Lee, S. K., Yoon, S. Y., & Won, J. S., 2018: Vegetation height estimate in rice fields using single polarization TanDEM-X science phase data. *Remote Sensing* 10(11), 1702.

Li, N., Lopez-Sanchez, J. M., Fu, H., Zhu, J., Han, W., Xie, Q., Hu, J., Xie, Y., 2022: Rice Crop Height Inversion from TanDEM-X PolInSAR Data Using the RVoG Model Combined with the Logistic Growth Equation. *Remote Sensing* 14(20), 5109.

Liao, C., Wang, J., Shang, J., Huang, X., Liu J. & Huffman T., 2018. Sensitivity study of Radarsat-2 polarimetric SAR to crop height and fractional vegetation cover of corn and wheat, *International Journal of Remote Sensing*, 39:5, 1475-1490.

Meier, U., 2018. Growth stages of mono- and dicotyledonous plants: BBCH Monograph doi:10.5073/20180906-074619

Mercier A., Betbeder J., Baundry J., Le Roux V., Spicher F., Lacoux J., Roger D., Hubert-Moy L., 2020. Evaluation of Sentinel-1 & 2 time series for predicting wheat and rapeseed phenological stages, *ISPRS Journal of Photogrammetry and Remote Sensing* 163, 231–256.

MGM 2023, Meteoroloji Genel Mudurlugu, https://www.mgm.gov.tr/veridegerlendirme/il-ve-ilceleristatistik.aspx?k=undefined&m=ANKARA (Dec 2023)

Nasirzadehdizaji, R., Balik Sanli, F., Abdikan, S., Cakir, Z., Sekertekin, A., Ustuner, M., 2019: Sensitivity analysis of multitemporal Sentinel-1 SAR parameters to crop height and canopy coverage. *Applied Sciences* 9(4), 655.

Ndikumana, E., Ho Tong Minh, D., Dang Nguyen, H. T., Baghdadi, N., Courault, D., Hossard, L., & El Moussawi, I., 2018: Estimation of rice height and biomass using multitemporal SAR Sentinel-1 for Camargue, Southern France. Remote Sensing, 10(9), 1394.

R Core Team 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. J. O. G. R. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geology Reviews, 71, 804-818.

Romero-Puig, N., Lopez-Sanchez, J. M., 2021: A review of crop height retrieval using InSAR strategies: Techniques and challenges. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14, 7911-7930.

Xie, Q., Wang, J., Lopez-Sanchez, J. M., Peng, X., Liao, C., Shang, J., Zhu, J., Fu, H., Ballester-Berman, J. D., 2021: Crop height estimation of corn from multi-year RADARSAT-2 polarimetric observables using machine learning. *Remote Sensing* 13(3), 392.

SNAP Development Team, 2023. Sentinel Application Platform Software, http://step.esa.int/main/toolboxes/snap/ (5 December 2023).

Yang, H., Li, H., Wang, W., Li, N., Zhao, J., & Pan, B., 2022: Spatio-temporal estimation of rice height using time series Sentinel-1 images. *Remote Sensing*, 14(3), 546. Yılmaz E. & Çiçek I., 2018. Detailed Köppen-Geiger climate regions of Turkey, International Journal of Human Sciences, 15, 1, 225-242.

Yuzugullu, O., Erten, E., & Hajnsek, I., 2018: Assessment of paddy rice height: Sequential inversion of coherent and incoherent models. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11(9), 3001-3013.